ISO/IEC JTC1/SC2/WG2 N2948R
L2/05-159R
2005-08-01

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации

Doc Type: **Working Group Document**
Title: **Proposal to add the Vai script to the BMP of the UCS**
Source: **Michael Everson, Charles Riley, and José Rivera**
Status: **Individual contribution**
Date: **2005-08-01**

Introduction

While there are divergent opinions on the ultimate origins of the Vai script and early influences on its development, there is general agreement that its beginnings as an actively used script are appropriately traced to Mɔmɔlu Duwalu Bukɛlɛ of Jondu, in what is now Grand Cape Mount County, Liberia. He is regarded within the Vai community, as well as by most scholars, as the script's inventor and chief promoter when it was first documented in the 1830s.

One of Bukɛlɛ's cousins, Kaali Bala Ndolɛ Wano, took as his "book-name" the *nom de plume* Rora, and produced a manuscript text of about 50 pages, *The Book of Ndole* (henceforth *Ndole*, also known as *The Book of Rora*), translated and analyzed at different times by Sigismund Koelle, Heymann Steinthal, and Gail Stewart. Other texts were written in the Vai script continuously during each decade of the 19th century, but the archives held at Jondu and Bandakɔro were destroyed in warfare with the neighboring Gola tribe. Those texts still extant from this period are of much shorter length than *Ndole*, including sample manuscripts obtained separately by the American missionaries J. L. Wilson and S. R. Wynkoop in 1834, and by F. E. Forbes in 1849.

Koelle's pioneering work on Vai began with a trip into Vai country to meet Bukɛlɛ in 1849. He compiled an inventory of the characters, including several logograms then in use, and translated three Vai books. He continued referring to his work on Vai into his later years, including notes and comments published in 1884 and 1889. Other researchers who took a particular interest in Vai, following Koelle, included Maurice Delafosse, F. W. Migeod, George Washington Ellis, and August Klingenheben; and in its more recent period (post-1962), P. E. H. Hair, David Dalby, Svend Holsoe, Sylvia Scribner, Michael Cole, John Singler, and Konrad Tuchscherer.

At the beginning of the 20th century, Momolu Massaquoi had undertaken efforts to translate gospel and Qur'anic materials into Vai, proposing modifications of characters for use in representing foreign sounds. Becoming Consul to Hamburg and beginning a teaching post at the University of Hamburg, he was able to continue this work in close collaboration with Klingenheben through the 1920s and 1930s. Klingenheben's involvement culminated with a conference held at the University of Liberia in 1962 to standardize the script for modern usage; the team of Vai scholars he worked with included Fatima Massaquoi Fahnbulleh, Zuke Kandakai, S. Jangaba Johnson, and Bai Tamia Moore, among others.

Texts produced over the course of the 20th century included a manuscript of 180 pages in length in diary form, containing a number of clan histories, translations of selections from the gospels and the Qur'an, folktales and short stories, and page-long summaries included in local newsletters. This script remains in use today, particularly among Vai merchants and traders. In addition to its presence in commerce, there is a growing body of literature published in Vai. The

Bible Society in Liberia has recently published a New Testament and the Institute of Liberian Languages has published several compilations of folktales and history.

Sources

The primary sources for the Vai characters in the character set proposed are the 1962 Vai Standard Syllabary (which was a distillation of many sources specifying characters for modern use), modern primers and texts which use the Standard Syllabary (and a few glyph modifications reflecting modern preferences), the 1911 additions of Momolu Massaquoi, and the characters found in *The Book of Ndole*. Secondary sources, such as Johnston 1906 and Dalby 1967, are used as supplementary material and as checks for some of the archaic characters. Both of those sources make some unifications and some distinctions which we have not considered authoritative in terms of the proposed character set. Only those characters in Dalby's "unidentified" column which are circled in Figure 8 do we consider to be safely identified by us. Many of the characters on Dalby's charts are not proposed for encoding.

Structure

In Figure 1 the syllabary as proposed here is given in its standard presentation. Vai is a simple syllabic script written from left to right. (Strictly speaking, the writing system is based on the mora, a unit of duration [or weight] such that a short syllable has one mora and a long syllable has two. A syllable is long if it contains a long vowel or ends with a consonant) The Vai language has seven oral vowels [e i a o u ɔ ɛ] and five nasal vowels [ĩ ã ũ ɔ̃ ɛ̃]. Vai has 31 consonants [ŋ h w p b ɓ mɓ kp mgb gb f v t d l r ɗ nɗ s ʃ z tʃ dʒ ndʒ j k ŋg g m n ɲ] of which [r] and [ʃ] are recent imports into the language. The original Vai script accounted for only the original 29 consonants, but by 1911 Massaquoi's charts give an extended set of characters, including [r ʃ] and also giving series for [ʍ θ ð ʒ]. All six of these extensions were modifications of the base glyphs for [l s w t d z]. Massaquoi's charts also fill in some of the "gaps" for syllables with forms for nasal close vowels [ẽ õ]. His series for [ʍ] was taken over in modern times to represent the nasal [w] series, and his series for [r] has been adopted as part of the Standard syllabary. Massaquoi's other additions for non-Vai sounds have been shown there in red in order to highlight them to the Vai-speaking reviewer of this document who may not be familiar with them. Massaquoi's charts are given in Figure 4.

Diacritical marks

Vai does not make use of a productive system of diacritical marks. Although the glyphs for some characters are certainly related to those of other characters, no systematic application of diacritization is found in Vai, and each character is encoded uniquely.

Vowel length

Modern Vai marks vowel length in three traditional ways. The most common orthographic practice is to write the length by echoing the vowel of the syllable with a syllable in the 𖬲 HEE series; so for 𖬁 FA with a long vowel, 𖬁𖬰 FAHA *faa* is written; for 𖬇 NI with a long vowel, 𖬇𖬱 NIHI *nii* is written. Alternatively, but rarely, the 𖬵 SYLLABLE LENGTHENER can be used: 𖬁𖬵 FA: *faa*, 𖬇𖬵 NI: *nii*. For the third method, namely, the "inherent vowel length" found in some Vai symbols, see see the section "Logograms" below.

Punctuation

Vai makes use of European punctuation, though a small number of script-specific punctuation marks occur commonly. The ˄ COMMA rests on or slightly below the baseline; the * FULL STOP rests on the baseline and can be ** doubled for use as an exclamation mark. The 𖬹 QUESTION

MARK also rests on the baseline; it is rather rarely used. Some modern primers prefer Vai punctuation; some prefer European. Vai punctuation was used in the 1962 Standard.

Digits

In the 1920s ten decimal digits were devised for Vai; these were "Vai-style" glyph variants of European digits (see Figure 11). They were not popular with Vai people and are not proposed for encoding here, even for historical purposes, as we have seen no examples of them in use. All the modern literature uses European digits.

Syllables unique to *The Book of Ndole*

In some of the older Vai texts more than one character may be used to represent a pronounced syllable; where these coexist beside another attested character, a disunification is warranted. We mention glyph choices in the discussion below because this is the first time many of these symbols have been given modern typographic form. To design modern typographic forms for such characters, we have looked first at characters in Ndole which resemble them, and then at how those glyphs compare to their modern counterparts; glyphs for the archaic characters were designed on the same principles. The same is true for the design of the logograms discussed below.

- ꕀ/ꕁ In *Ndole* the character ꕀ is read as a logogram KAI 'man', but this is the same character as the modern ꕀ SYLLABLE KA, and in *Ndole* ꕀ and ꕁ SYLLABLE NDOLE KA are distinct; compare the related ꕂ SYLLABLE GA. Dalby does not give the ꕁ shape.
- ꕃ/ꕄ In *Ndole* ꕃ SYLLABLE SOO is distinct from the ꕄ SYLLABLE NDOLE SOO (Stewart suggested that the former was originally *SOONG). The glyph we give for the latter reflects the same glyph changes as shown in Dalby's 1849 and 1961 glyphs for ꔅ SYLLABLE PE, namely reversal and sharpening of the centre curve into a zigzag.
- ꘋ Stewart says that there are four "variant" characters used only in names in *The Book of Ndole*. Of these, the ꘋ SYLLABLE NDOLE FA is certainly a candidate for encoding, distinct from both ꔆ SYLLABLE FA and ꖜ SYMBOL FAA 'die, kill'. For the modern glyph we have used the shapes of ꖜ SYMBOL FAA to give it typographic form, based on the relationship of the Ndole FA and FAA glyphs; another choice might be to use a permutation of ꔇ SYLLABLE MBA, which also has similar shapes. Of the other "variant" characters, two of them are already encoded for other uses: ꘈ SYLLABLE SHA and ꘎ SYLLABLE WOON are available to represent the *Ndole* "variants" SA and WO respectively. Further research into the "variant" MA Stewart describes will be required to see if it really should be disunified from ꔐ SYLLABLE MA; to us, it seems like a unifiable glyph variant. It could be added at a later date.

Logograms

In the oldest Vai texts, a set of symbols called in the literature "logograms" occurs, representing individual syllables with inherent vowel length or with an inherent final nasal. Two of these logograms are known to enjoy some contemporary use. Of these, the character ꖜ SYMBOL FAA, has an inherent long vowel. It represents the word meaning 'die, kill' and is used alongside the date of death, as the † DAGGER is in some European practices (the glyph is said to represent a wilting tree). The word meaning 'thing' is represented by ꖆ SYMBOL FEENG. A list of the logograms occurring in *Ndole* is given in Figure 6; a number of these are also represented by Dalby in Figure 8. Examination of these repertoires shows that some of the logograms were absorbed as ordinary syllables in the modern syllabary; others we judge to be distinct and have proposed here.

𞄀 In *Ndole* 𞄀 SYMBOL BANG, 𞄁 SYLLABLE BA, and 𞄂 SYLLABLE BHA are distinct. Dalby seems to mistakenly identify 𞄀 SYMBOL BANG with 𞄂 SYLLABLE BHA, listing both under **ba** (BHA). The glyph we give for 𞄀 SYMBOL BANG reflects the same glyph changes as shown in Dalby's 1849 and 1961 glyphs for both 𞄃 SYLLABLE DEE and 𞄄 SYLLABLE SU, namely rotation and sharpening of the hooks into points.

𞄅 Both *Ndole* and Dalby give 𞄅 SYMBOL DANG 'hear, understand'. For the glyph shape we give, compare Dalby's 1849 and 1961 𞄆 SYLLABLE MI.

𞄃 Although the glyph Dalby gives for 1841 𞄃 SYLLABLE DEE has dots or rings, and *Ndole* SYMBOL DEENG 'child, small' has not, we believe these characters should be unified. *Ndole* gives no distinct SYLLABLE DEE, and rotation and sharpening of the hooks gives the modern glyph shape 𞄃; compare Dalby's 1849 and 1961 glyphs for 𞄄 SYLLABLE SU. Johnston 1906 also gives the character, with dots or rings, with the reading *de* (see Figure 17).

𞄇 We consider it possible but not necessarily probable that Dalby's 1849 **dòn·** (DHOONG) is the same character as the *Ndole* 𞄇 SYMBOL DOONG 'enter'. If a *SYMBOL DHOONG exists, we do not know it apart from Dalby's suggestion, and have not encoded it separately here. In any case for the glyph for SYMBOL DOONG we have followed that for 𞄈 SYLLABLE NYA in modernizing it from the *Ndole* manuscript.

𞄉 *Ndole* gives 𞄉 as a logogram for SYMBOL DO-O 'be small' [ɗɔː] (not DOO [ɗo]), although Dalby suggests that it is an 1849 glyph variant of modern 𞄊 SYLLABLE NDOO; we note that Dalby gives another SYMBOL DHUNG which looks similar to NDOO, and we note the general similarity of 𞄋 SYLLABLE LOO, not to mention Massaquoi's 𞄋 SYLLABLE THO. We think on balance Dalby's unification is unsafe, and consider it wiser not to try to unify SYMBOL DO-O 'be small' with SYLLABLE NDOO. *Ndole* does not otherwise give a SYLLABLE NDOO, but both its SYLLABLE LOO and SYLLABLE DOO have the three branches at the top and a loop below. Johnston 1906 also gives the character with the reading *dō* (dɔ/ɗɔ). See Figure 17.

𞄌 Both *Ndole* and Dalby give 𞄌 SYMBOL FAA 'die, kill', though Dalby gives it as a glyph variant of 𞄍 SYLLABLE FA. The glyph shape we give conforms to modern practice.

𞄎 Both *Ndole* and Dalby give 𞄎 SYMBOL FEENG 'thing'. The glyph shape we give conforms to both older and modern practice.

𞄏 Both *Ndole* and Dalby give 𞄏 SYMBOL JONG 'slave'. The glyph shape follows *Ndole*, with influence from 𞄐 SYLLABLE JEE.

𞄑 Dalby reads 𞄑 as 𞄒 SYLLABLE GI but marks his reading as doubtful. We share that doubt and propose to encode *Ndole*'s 𞄑 SYMBOL KEENG 'foot' as a unique character. Johnston 1906 also gives this character with the reading *keñ* (KEENG). See Figure 18.

𞄓 In *Ndole* 𞄓 SYMBOL NII 'cow' is given, alongside a second reading KPE KOWU 'case of gin'. Johnston 1906 also gives this character with the reading *ni* (see Figure 19). We find Dalby's unification of this character with 𞄔 SYLLABLE NYU to be doubtful in this context, and prefer to encode the symbol on its own. Neither *Ndole* nor Johnston 1906 give a syllable NYU, as it happens, but even so, since [n] and [ɲ] are different phonemes and the internal structure fo the script shows no relation between glyphs of those two phonemes, we think it unlikely that the characters are related. Other early box-like glyphs retain a box-shape and do not take on a U-shape; Massaquoi's glyph for NYU has but a single dot (see Figure 4).

𞄕 Both *Ndole* and Dalby give 𞄕 SYMBOL KUNG 'head, be able'. The glyph shape follows Dalby, as this harmonizes better with modern 𞄖 SYLLABLE KPA than a tilted triangle would.

- The glyph Dalby gives for 1849 SYLLABLE DO is the modern glyph turned on its side; we believe that *Ndole* SYMBOL LO 'in' is the same character and the two should be unified. *Ndole* seems to have some confusion using SYLLABLE LO for LO as well as DO and NDO; it also uses for LA, DA, NDA; for LA, DA; for LU, DU; for LE , DE; and for LOO, DOO. Johnston 1906 likewise shows early confusion in the representation of these phonemes: for RA, DA, NDA; for RI, DI; for ROO, DOO; and for RO , DO (Johnston does not use L- in transcription).
- Both *Ndole* and Dalby give SYMBOL TAA 'go, carry, journey'. The glyph shape we give conforms to both older and modern practice.
- Both *Ndole* and Dalby give SYMBOL TING 'island'. The glyph shape we give conforms to both older and modern practice.
- Both *Ndole* and Dalby give SYMBOL TONG 'be named'. The glyph shape we give conforms to both older and modern practice.

Character names

Transliteration into ASCII-based character names follows the usual UCS conventions. The vowels are EE I A OO U O E, doubling the close vowels [e o] and leaving single the open vowels [ɔ ɛ]. Nasality is marked with -N. Because the implosive consonants are far more frequent than the non-implosive ones, Vai are [ɓ ɗ] conventionally transliterated <b d> and [b d] are conventionally written <bh dh>; accordingly [ð], a modification of <dh> is written <dhh> here. Figure 1 gives both phonetic values and UCS character name values.

Ordering

There is no evidence of traditional conventions on ordering *per se* apart from conventions found in syllabary charts. The ordering proposed here is that agreed upon by an experts working group, based on current practices in chart presentation. Our analysis of Vai recognizes that structurally, the "inherent" order is based first on the rhyme, and then on the consonant-groupings within each rhyme, themselves in a standard order which shows a high degree of linguistic sophistication on the part of the script's creator, Mɔmɔlu Duwalu Bukɛlɛ.

Consonants. The structure of the Vai glyphs as traditionally given in syllabary charts shows a vertical glyph relation between many characters; accordingly, a visual sort which preserves this relationship makes sense for assisting readers in finding characters in lists. In this "column-based" sort, for each rhyme, the full run of consonants from Ø to *ny-* is given for the [e] vowel, then the next column of consonants from Ø to *ny-* is given for the [i] vowel, and so on to the [ɛ] vowel. The relative order of the consonants is based on the relationships of the glyphs in the script itself. This ensures that similar letters would appear near each other in the alphabet. according to the rhymes first, then the similarly-shaped letters are next to each other: > > >

Latin transliteration of Vai characters is found; others are "column-based" as though with the vowel orders sorted as in Latin [a ɛ e i ɔ o u] or [a e ɛ i o ɔ u]; a "linguists' order" [i a u ɛ e o ɔ] is sometimes found, as in Dalby 1967.

Logograms. The NDOLE SYLLABLES 𖬁 FA, 𖬂 KA, and 𖬃 SOO are sorted as equivalent to the SYLLABLES 𖬁 FA, 𖬂 KA, and 𖬃 SOO. The SYMBOLS 𖬄 FEENG, 𖬅 KEENG, TING, 𖬆 BANG, 𖬇 DANG, 𖬈 DOONG, 𖬉 KUNG, 𖬊 TONG, and 𖬋 JONG are sorted as equivalent to the SYLLABLES 𖬍 FEE+NG, 𖬎 KEE+NG, 𖬏 TI-NG, 𖬐 BA-NG, 𖬑 DA-NG, 𖬒 DOO-NG, 𖬓 KU-NG, 𖬔 TO-NG, and 𖬕 JO-NG respectively. The SYMBOLS 𖬖 NII, 𖬗 FAA, 𖬘 TAA, and 𖬙 DO-O are sorted as equivalent to the SYLLABLES 𖬛 NI+LENGTHENER, 𖬗 FA+LENGTHENER, 𖬘 TA+LENGTHENER, and 𖬚 DO+LENGTHENER. We do not recommend that these be treated as equivalent to the SYLLABLES 𖬜 NI+HI, 𖬝 FA+HA, 𖬞 TA+HA, and 𖬟 DO+HO, even though the readings are the same.

Line breaking

An opportunity for line breaking within a word can occur *after* any character, but not *before* U+A606 𖬈 SYLLABLE NG or U+A606 𖬛 SYLLABLE LENGTHENER which should not begin a line in hyphenation since they are syllable finals. (𖬈 SYLLABLE NG can begin a line when it is the first-person singular pronoun, but then it does not follow another character.) The "logogram" symbols should be treated as any other character. In words where length is indicated by a 𖬛 HEE series character echoing the vowel of the preceding syllable, the long syllable should not be broken. While VA + I = 𖬡 *Vai* can be broken, VA + HA 𖬡𖬞 *vaa* should not be broken. The punctuation marks U+A608 VAI COMMA, U+A609 VAI FULL STOP, and U+A60A VAI QUESTION MARK can only have a break after (but when VAI FULL STOP is used twice as an.exclamation mark the pair should not be broken).

Unicode Character Properties

All Vai syllables and symbols have the same properties except for the three punctuation characters:

```
A500;VAI SYLLABLE EE;Lo;0;L;;;;;N;;;;;
..
A605;VAI SYLLABLE NYE;Lo;0;L;;;;;N;;;;;
A606;VAI SYLLABLE NG;Lo;0;L;;;;;N;;;;;
A607;VAI SYLLABLE LENGTHENER;Lo;0;L;;;;;N;;;;;
A608;VAI COMMA;Po;0;CS;;;;;N;;;;;
A609;VAI FULL STOP;Po;0;CS;;;;;N;;;;;
A60A;VAI QUESTION MARK;Po;0;ON;;;;;N;;;;;
A60E;VAI SYMBOL FENG;Lo;0;L;;;;;N;;;;;
A60F;VAI SYMBOL FAA;Lo;0;L;;;;;N;;;;;
```

Bibliography

Dalby, David. 1967. "A survey of the indigenous scripts of Liberia and Sierra Leone: Vai, Mende, Loma, Kpelle and Bassa", in *African Language Studies* 8:1–51.

Kandakai, C. K., and Charles Hutchison. 1975. 𖬡𖬞 𖬀𖬟 𖬇𖬢𖬰 (*Vái kpòlò sénjé-nà*). Monrovia: Institute for Liberian Languages.

Jensen, Hans. 1969. *Die Schrift in Vergangenheit und Gegenwart*. 3., neubearbeitete und erweiterte Auflage. Berlin: VEB Deutscher Verlag der Wissenschaften.

Johnson, S. Jangaba M. [s.d.] *Traditional History, Customary Laws, Mores, Folkways, and Legends of the Vai Tribe*. Monrovia: Department of the Interior.

Johnston, H. 1906. *Liberia*. London: [s.n.].

Kandakai, Zuke, et al. 1962. 𖬡𖬞 𖬀𖬟 𖬌𖬞𖬦𖬢𖬣𖬤 𖬇𖬇 (*Vai kpolo saikilamaa mɛ*) = *The Standard Vai script*. Monrovia: University of Liberia African Studies Program.

Kante, Solomana. [s.d.]. ﻩﺎﻨﺪﻧﺎﻓ ﻮﺠﻨﻛ (*Manden kurufaba*). Edited by Baba Mamady Diane. [s.l.]: [s.n.].

Massaquoi, Momolu. 1911. “The Vai people and their syllabic writing”, in *Journal of the Royal African Society* 10.40, July, pp. 459-466.
Nyei, M. B., and H. Getaweh. 1976. 𖬏𖬵𖬰 𖬇𖬲𖬏 𖬒𖬏 𖬮 𖬑𖬵𖬵 𖬒𖬏 𖬮 𖬵𖬵𖬺𖬮𖬵 𖬯 (*Wodu Vai kpolo e poha kpolo e senje-na ã*). [s.l.]: [s.n.].
Scribner, Sylvia, and Michael Cole. 1981. *The Psychology of Literacy*. New York: [s.n.].
Stewart, Gail, and P. E. H. Hair. 1969. “A Bibliography of the Vai Language and Script”, in *Journal of West African Languages*, volume VI, number 2, p. 124.
Tucker, S. Lamii. 1999. 𖬇𖬲𖬏 𖬒𖬏 𖬓 𖬹𖬵𖬮 𖬵 𖬁𖬵𖬵 𖬆 𖬒𖬏𖬮 𖬊𖬺𖬵𖬯 (*Vai kpolo lɔ kowae gbu kpaha sɔ kpoloe fɔlanaã mɛ*) = *Vai script primer 2*. Monrovia: Libtralo.
Tuchscherer, Konrad, and P.E.H. Hair, “Cherokee and West Africa: Examining the Origins of the Vai Script”, in *History in Africa,* 29, 427-486.
[s.n.]. [s.d.]. 𖬲𖬵𖬷 𖬋𖬵𖬵 𖬇𖬲𖬏 𖬋 𖬮 𖬺 (*Iwaŋga kaha Vai wo e la*) = *Teach Yourself Vae: Book One*. [s.l.]: [s.n.].

Acknowledgements
This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley).

Figures

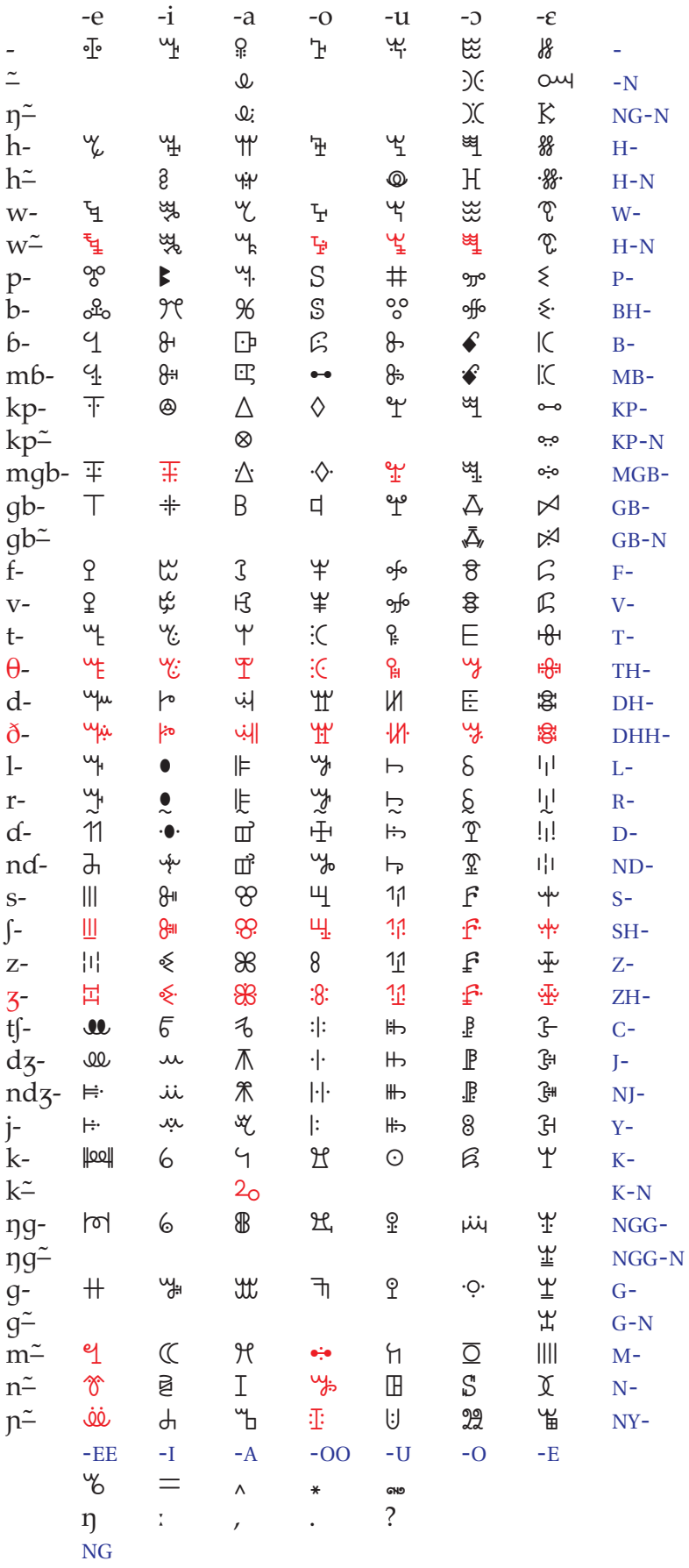


Figure 1. Vai character repertoire as conventionally presented in tabular form. The characters in red are Massaquoi's additions (apart from his *wh-* and *r-* series) as well as the syllable 𖬁𖬰 KAN from *The Book of Ndole*. On the left are phonetic values; in blue are the UCS name transcriptions.

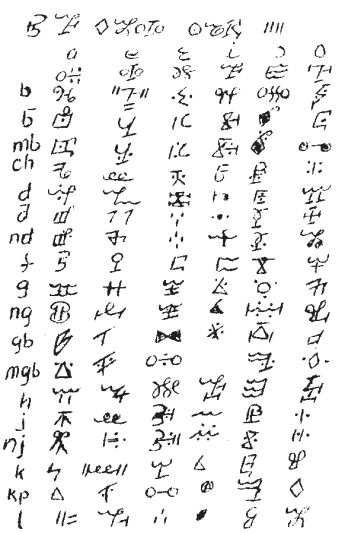


Figure 2. Handwritten syllabary from a teach-yourself book.

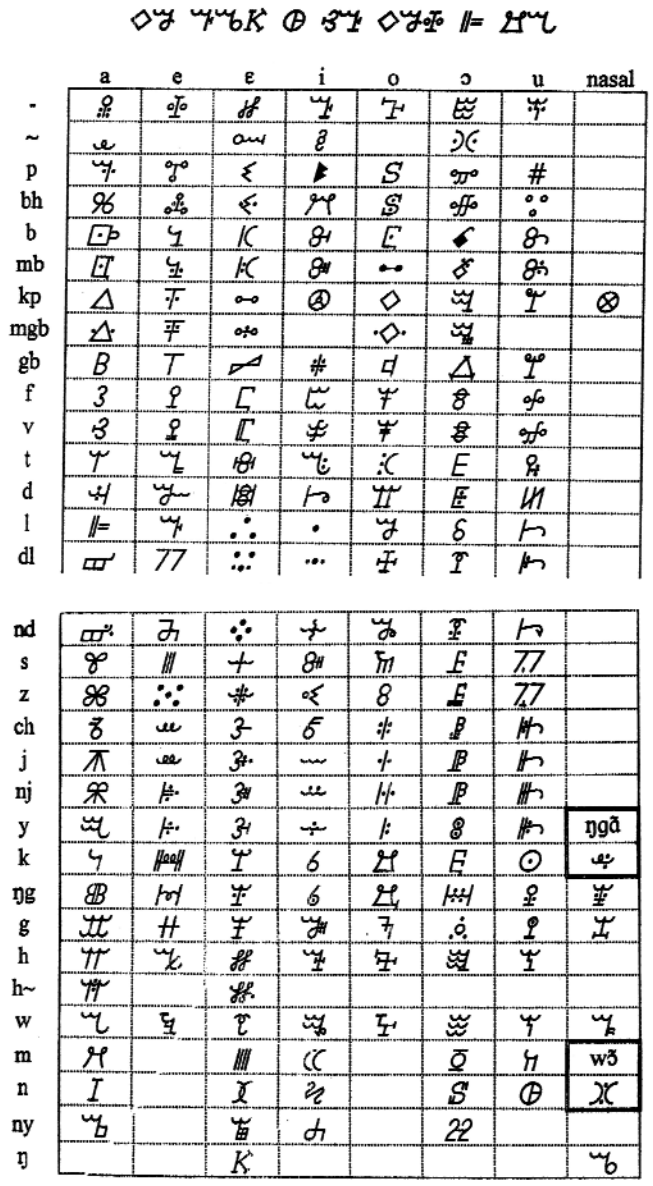


Figure 3. Vai syllabary chart from Tucker 1999.

Figure 4. Syllabary chart from Massaquoi 1911. The logogram FAA occurs in his "Miscellaneous" section; of his "Punctuation and other signs", only COMMA, QUESTION MARK, and FULL STOP have been proposed for encoding. As noted, the exclamation mark is simply two FULL STOPS.

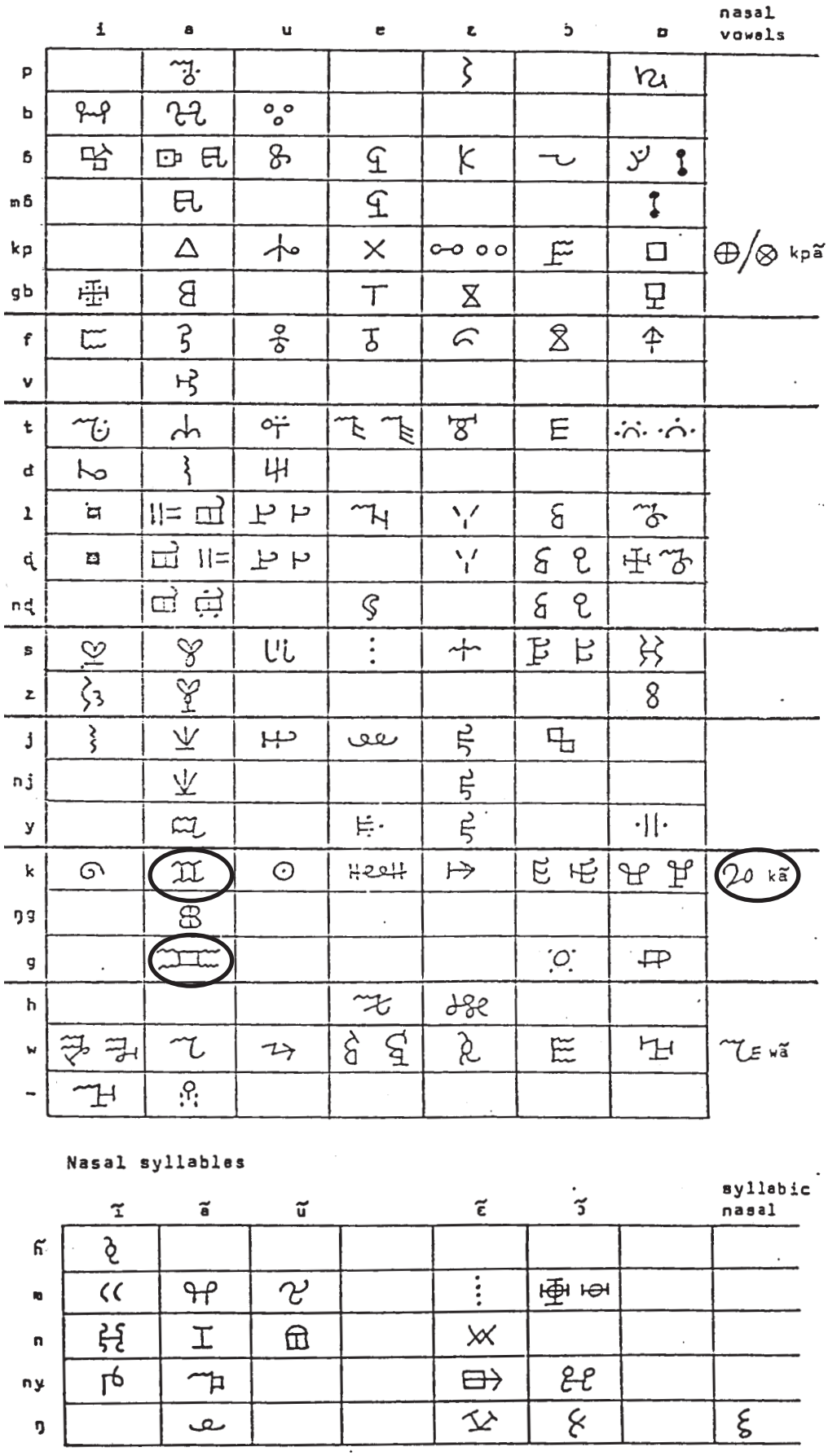


Figure 5. The "Ndole Syllabary", that is, the repertoire of character in *The Book of Ndole*. Circled is 𞤛 NDOLE SYLLABLE KA, not found in later syllable charts; note its similarity to 𞤜 SYLLABLE GA. Also circled is 𞤟 SYLLABLE KAN.

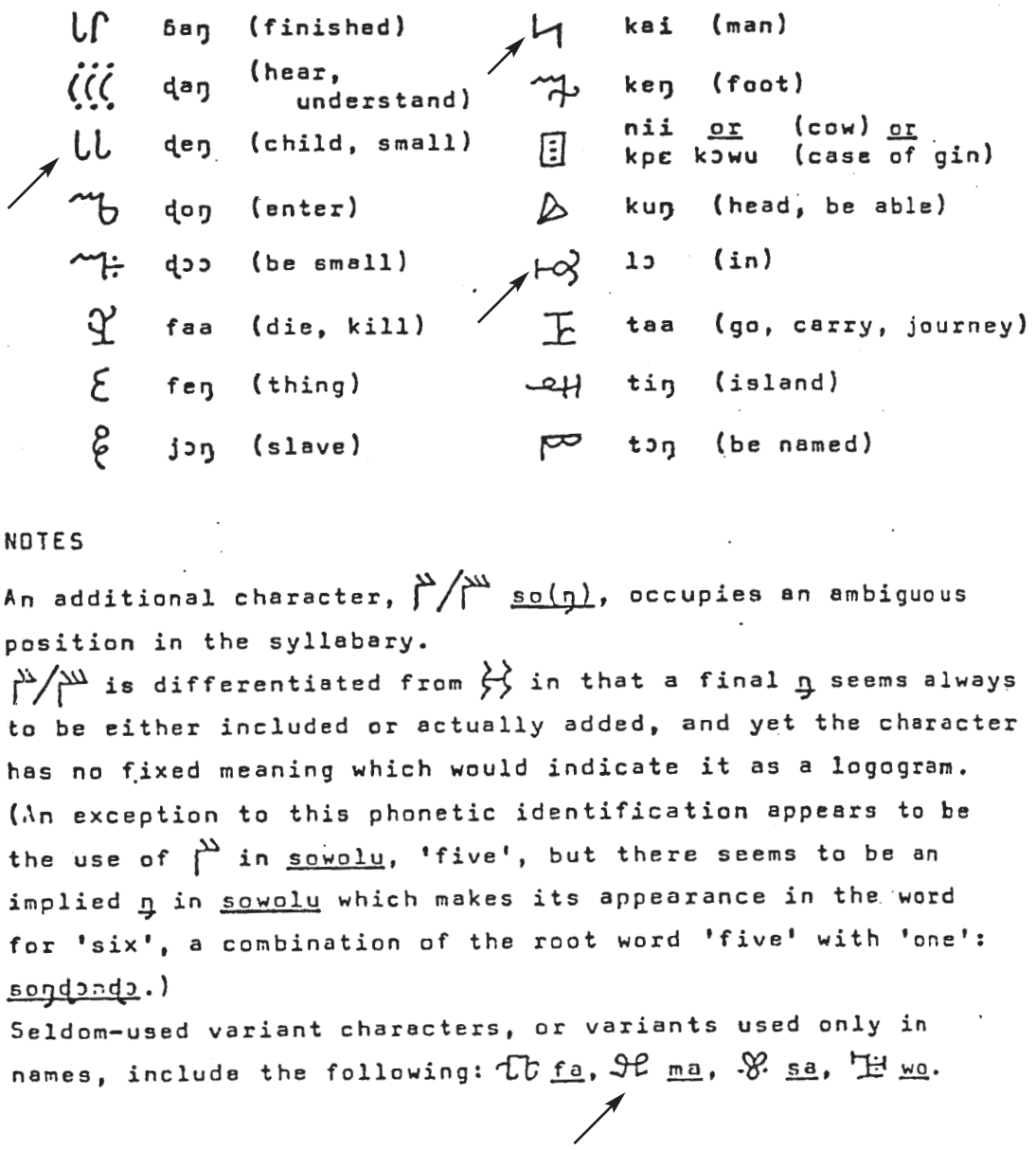


Figure 6. Logograms from *The Book of Ndole*. All but *SYMBOL DEENG, *SYMBOL KAI, *SYMBOL LO, and *SYLLABLE NDOLE MA are proposed for encoding, as they are considered to be unified with 𞤃 SYLLABLE DEE, 𞤊 SYLLABLE KA, 𞤐 SYLLABLE DO, and 𞤑 SYLLABLE MA.

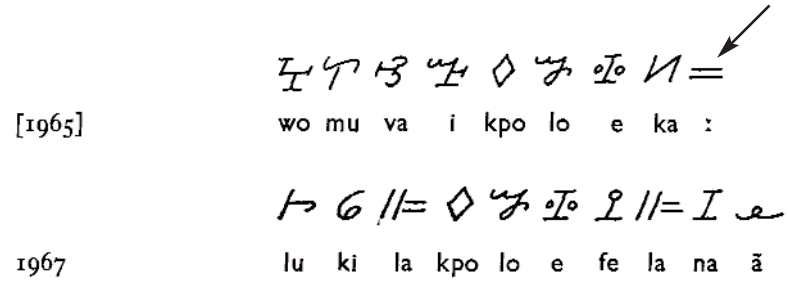


Figure 7. An example from Stewart and Hair 1969. showing the 𞥋 SYLLABLE LENGTHENER in use in a book title.

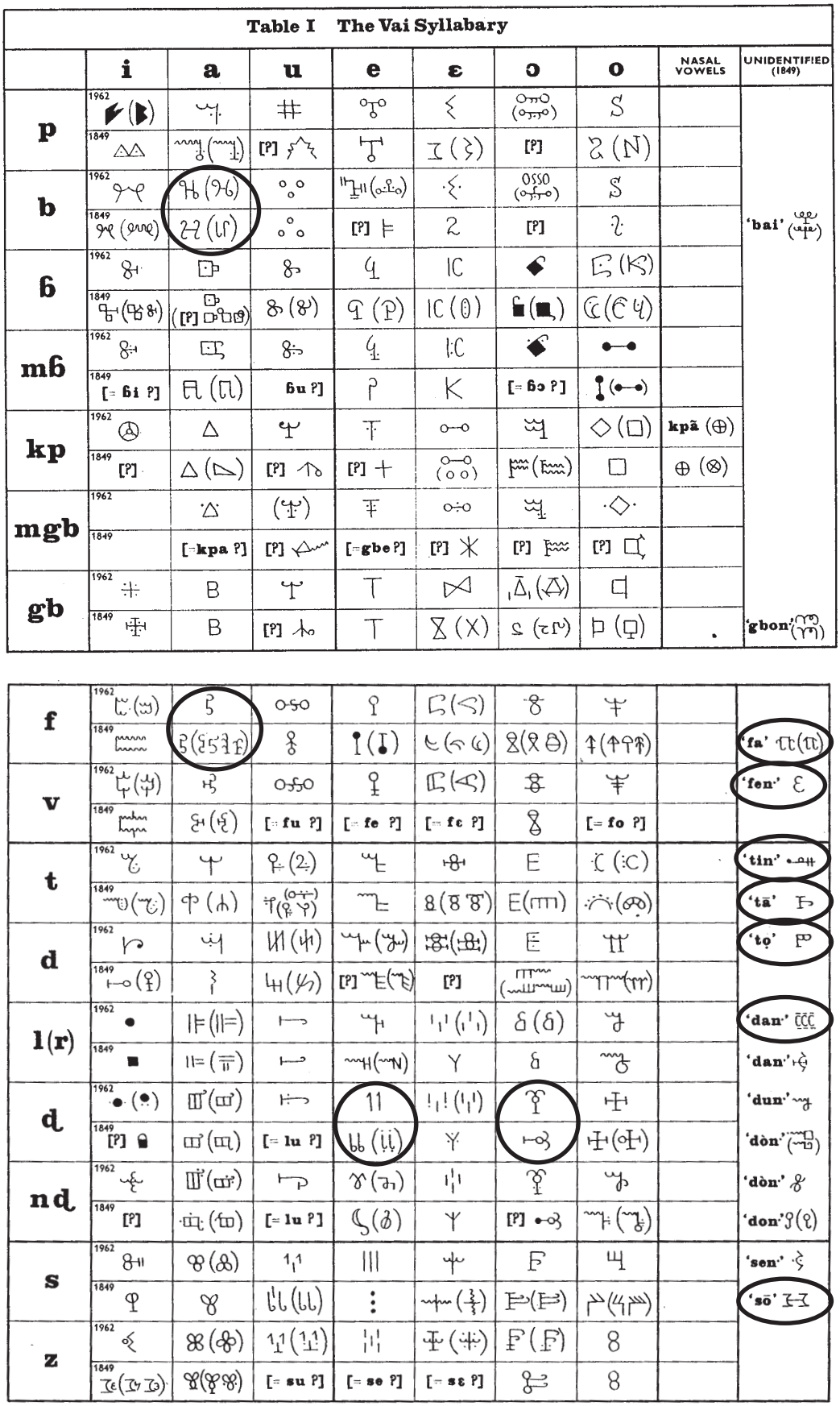


Figure 8a. Syllabary chart (here and on the next page) from Dalby 1967 showing old (1849) and contemporary (1961) syllable shapes. Circled here are SYMBOL BANG, SYMBOL FAA, SYLLABLE NDOLE FA, SYMBOL FEENG, SYMBOL TING, SYMBOL TA, SYMBOL TONG, SYMBOL DANG, SYMBOL DEE (used in *The Book of Ndole* for *SYMBOL DEENG), SYMBOL DO (used for *SYMBOL LO), and SYLLABLE NDOLE SOO.

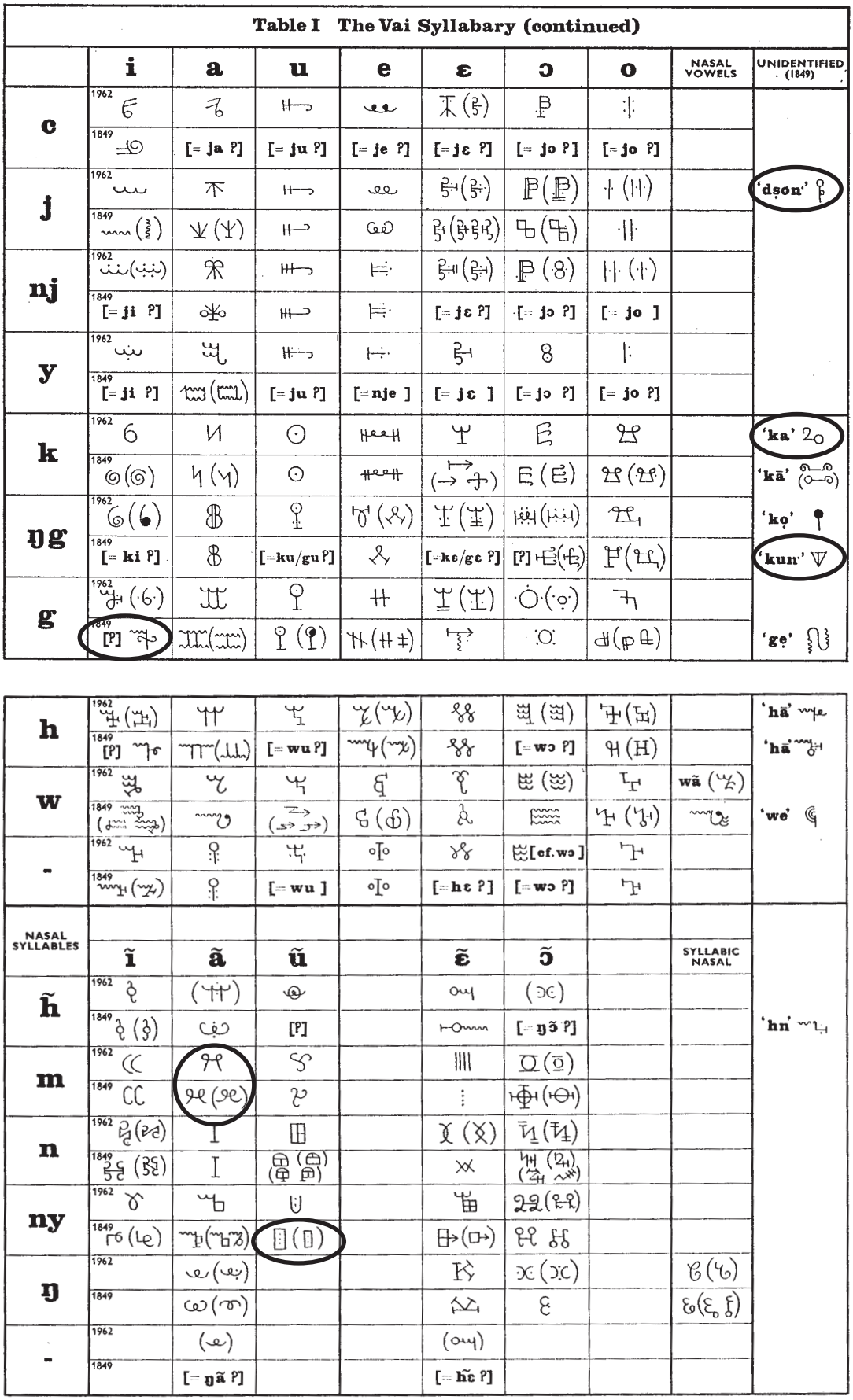


Figure 8b. Circled here are SYMBOL JONG, SYLLABLE NDOLE KA, and SYMBOL KUNG. Also circled is SYLLABLE MA, showing its glyph variant. Compare Stewart's mention of this in Figure 6 below.

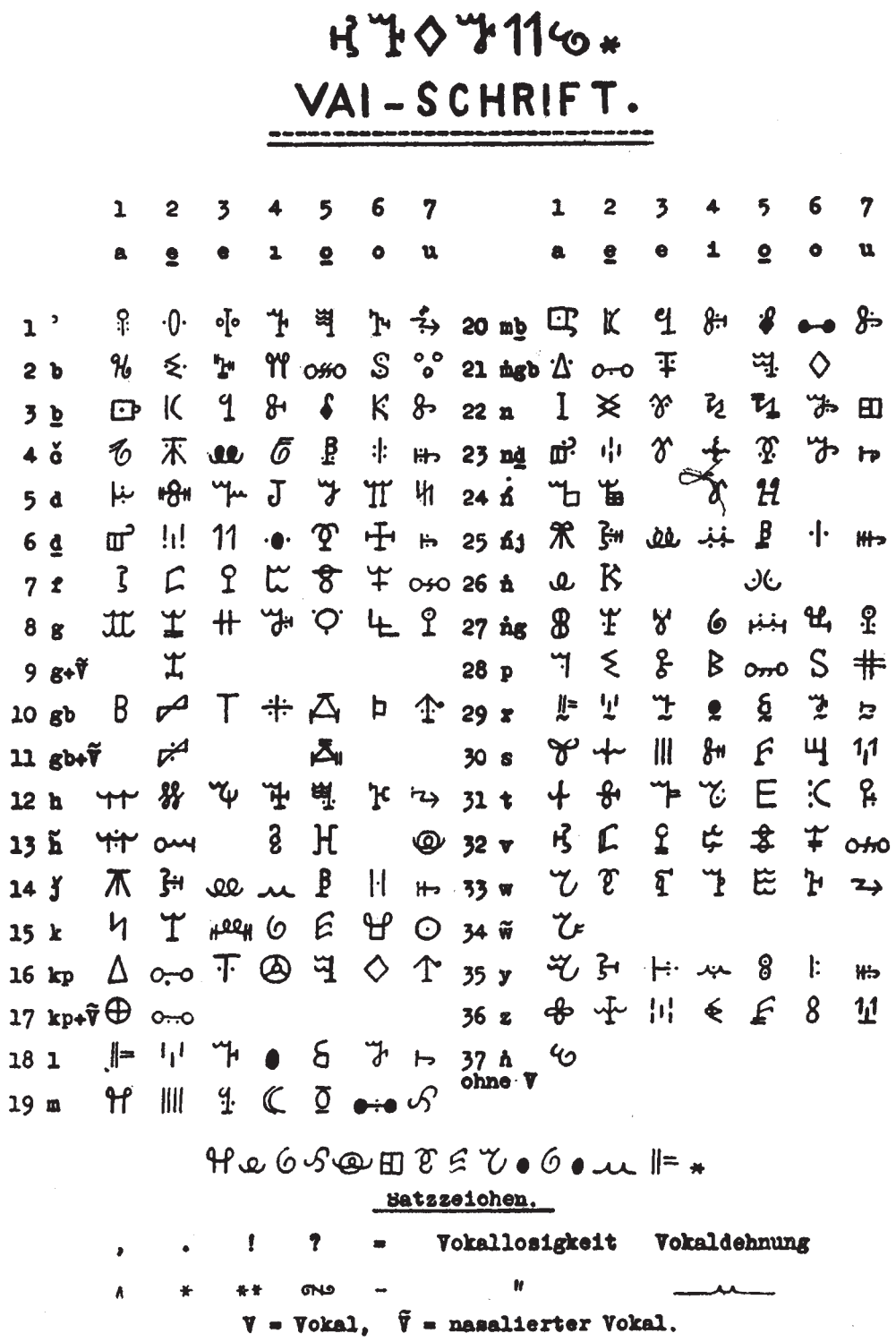


Figure 9. Syllabary chart entitled *Vai kpolo ɗeŋ* 'Vai script' from Jensen 1960. The "punctuation" symbols he gives for "vowellessness" (*Vokallosigkeit*) and "vowel length" (*Vokaldehnung*) seem to answer to what Massaquoi gives as "detraction" and "continuation of sound" in his chart in Figure 4.

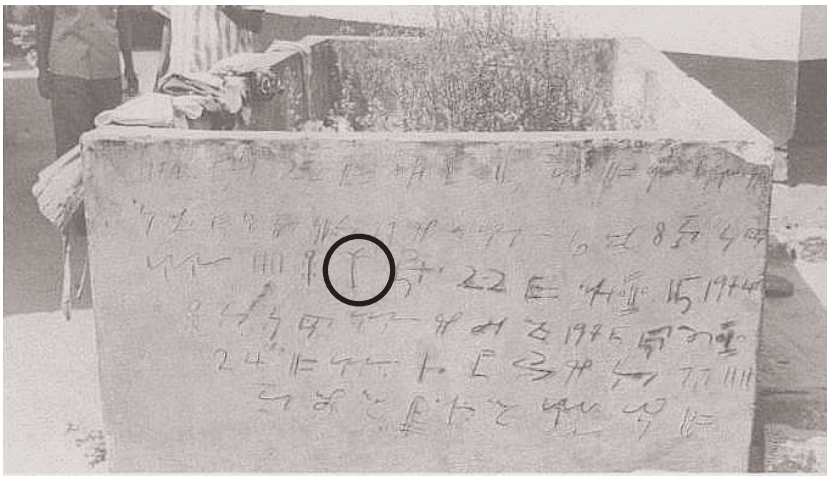


Figure 10. A tombstone in a Vai community in Monrovia, from Scribner and Cole 1981. The symbol 𖬁 FAA 'died' can be seen in the second line of the inscription.

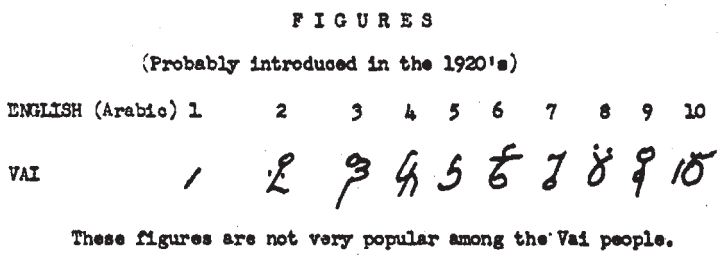


Figure 11. Vai digits, from S. Jangaba M. Johnson [s.d]. To date, this chart is the only representation of the Vai digits we have seen. They are not proposed for encoding because they are poorly attested; should they be needed in future, they can be encoded at that time.

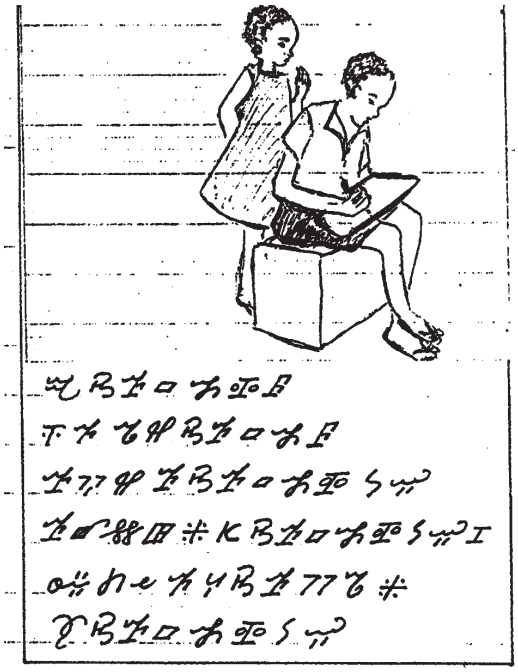


Figure 12. From Nyei and Getawek's 1976 schoolbook.

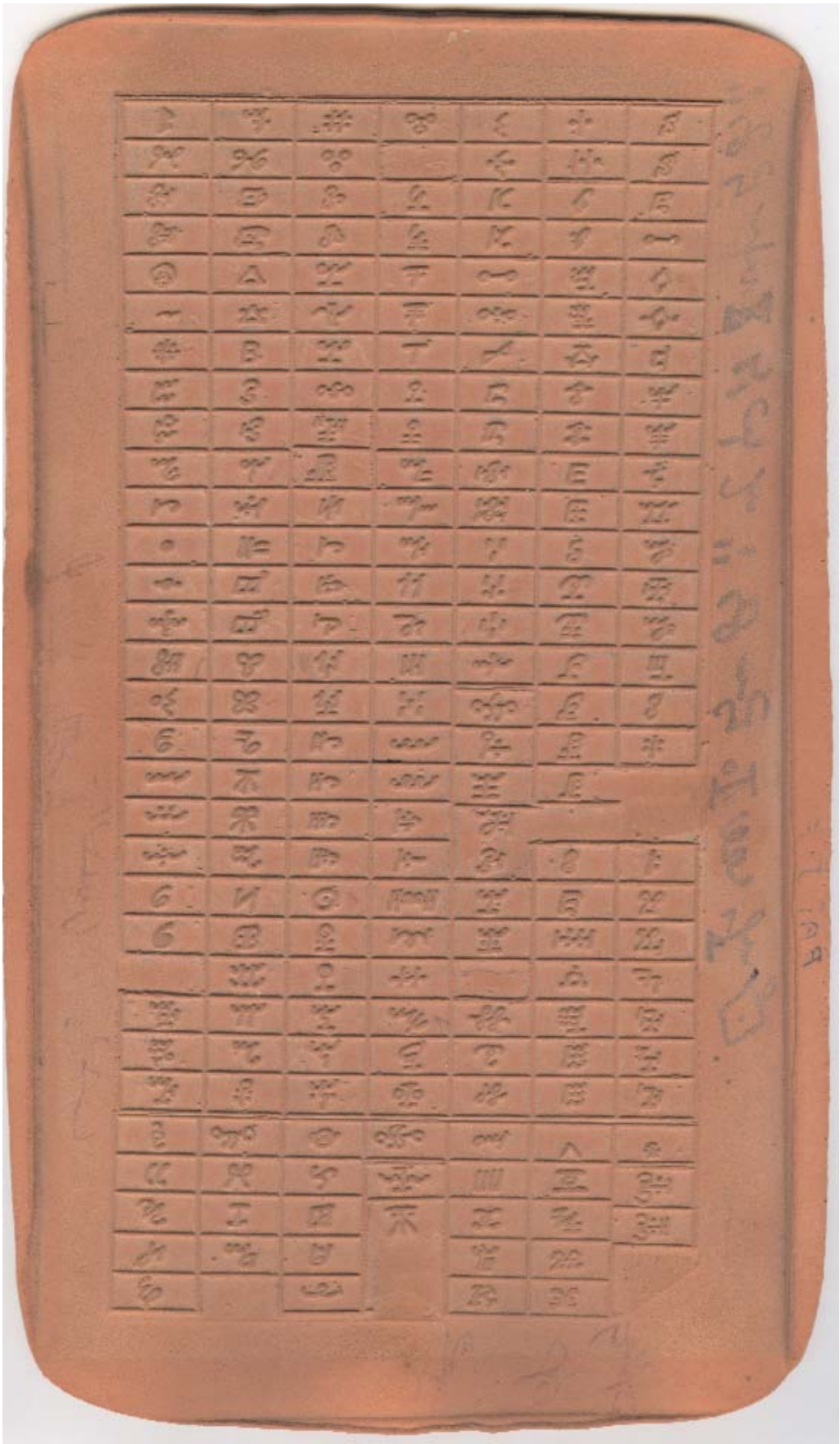


Figure 13. A rubber plate used to print the syllabary.
The image has been digitally reversed to show the characters as they would be printed.
The plate was produced by Scribner and Cole's 1981 project

Figure 14. Description of the Vai syllabary in N'Ko. by Solomana Kante.

Figure 15. Vai syllabary chart from Kandakai and Hutchison 1975.

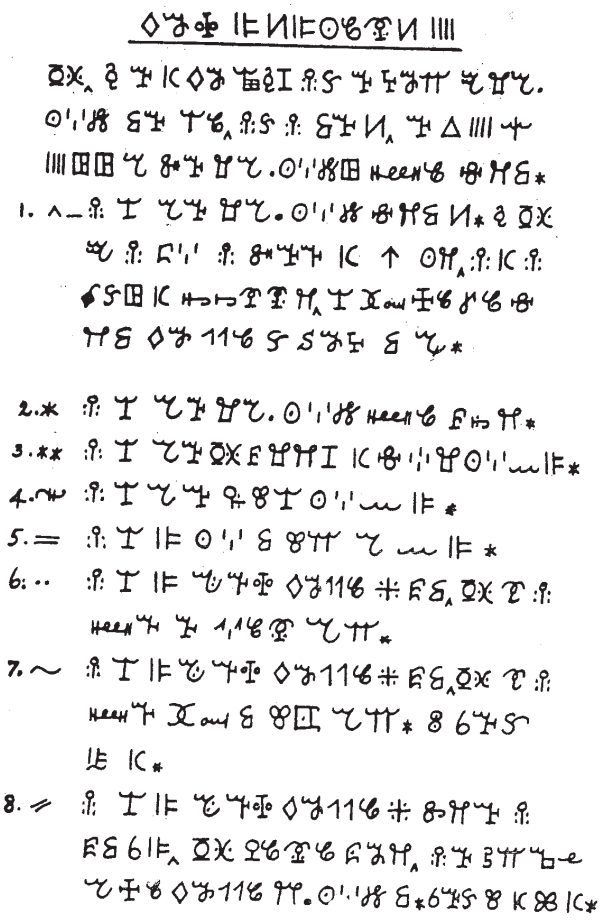


Figure 16. Discussion of Vai punctuation in Kandakai, et al., 1962. Items 1, 2, 3, 4, and 5 can be represented with the encoding proposed here. The second part of Item 2 is superceded by Massaquoi's series for [θ ð]. Items 6 and 8 do not seem to have ever been implemented at all (though item 6 first could have some utility, in keyboarding specifications); item 7 is not productive, and the *re*-series is fully encoded. This translation is by Mohamed Nyei:

This is information about reading the book

When you are writing and you want your ideas to be clear and precise, you should place these marks between your expressions.

1. ∧ This is what separates expressions. When one sees it written like this ↑ the associated character should be pronounced with the tongue between the teeth, as in English.
2. * This is what ends an expression or sentence.
3. ** This is used in commands to do something and denotes a statement of surprise.
4. ꘎ This denotes a question.
5. = This signifies a long sound.
6. .. When this is placed below a character, one should utter a nasal sound.
7. ~ When this is placed under a character, one should say it with a trilling tongue; like "ꕤ" *ra*.
8. // When this is placed beside a character to the right, it puts stress on the character; like in "ꕢ" *sa* and "ꕣ" *za*.

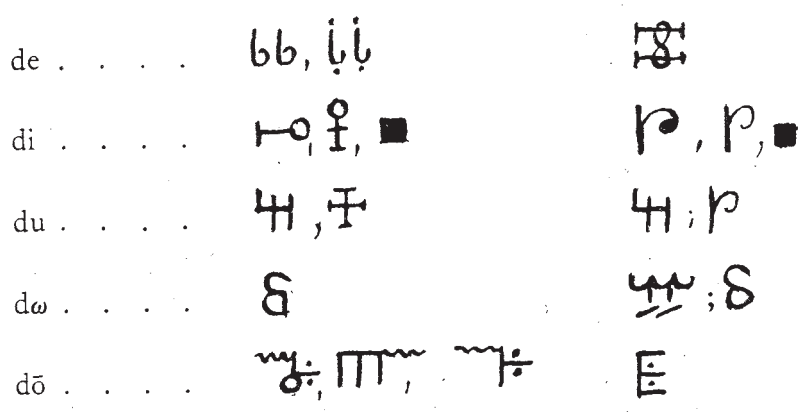


Figure 17. "Original" symbols on the left and "modern" symbols on the right in Johnston 1906. In the first line, Johnston shows *de* which corresponds to modern 𑖦 SYLLABLE DEE, though it has a ring shape or dot element. On the right the modern 𑖧 SYLLABLE DHE is shown; Johnston did not distinguish [d] from [ɗ] or [e] from [ɛ]. In the fourth line, Johnston shows *dō* with three glyphs, coresponding to what Dalby identifies as 1849 **nɗo** (NDOO), **dɔ** (DO), and **nɗo** (NDOO) respectively. We identify these instead as logogram 𖬲 ***dɔː** (SYMBOL DO-O), 𖬳 **dɔ** (DO), and 𖬴 **nɗo** (NDOO). On the right he shows the modern 𖬳 SYLLABLE DHO.

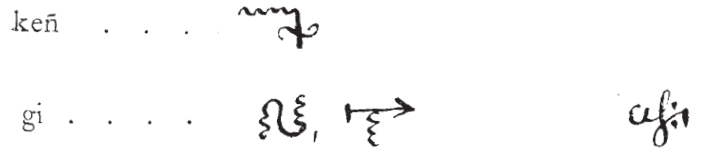


Figure 18. The syllable KEENG compared with the syllable GI from Johnston 1906. Dalby reads Johnston's glyph for *keñ* as ?**gi**, and gives the first two of Johnston's *gi* glyphs as "unidentified" **gẹ** and **gɛ** respectively. The modern glyph on the right is modern 𖬵 SYLLABLE GI. We have not found Johnston's first GI attested other than there and in Dalby; we have not proposed its encoding.

Figure 19. The syllable *ni* from Johnston 1906, showing both its 𑖩 SYLLABLE NI form and a 𖭃 SYMBOL NII form.

TABLE xx - Row A5: VAI

G = 00
P = 00

TABLE XX - Row A5: VAI

hex	Name
A500	VAI SYLLABLE EE
A501	VAI SYLLABLE HEE
A502	VAI SYLLABLE WEE
A503	VAI SYLLABLE WEEN
A504	VAI SYLLABLE PEE
A505	VAI SYLLABLE BHEE
A506	VAI SYLLABLE BEE
A507	VAI SYLLABLE MBEE
A508	VAI SYLLABLE KPEE
A509	VAI SYLLABLE MGBEE
A50A	VAI SYLLABLE GBEE
A50B	VAI SYLLABLE FEE
A50C	VAI SYLLABLE VEE
A50D	VAI SYLLABLE TEE
A50E	VAI SYLLABLE THEE
A50F	VAI SYLLABLE DHEE
A510	VAI SYLLABLE DHHEE
A511	VAI SYLLABLE LEE
A512	VAI SYLLABLE REE
A513	VAI SYLLABLE DEE
A514	VAI SYLLABLE NDEE
A515	VAI SYLLABLE SEE
A516	VAI SYLLABLE SHEE
A517	VAI SYLLABLE ZEE
A518	VAI SYLLABLE ZHEE
A519	VAI SYLLABLE CEE
A51A	VAI SYLLABLE JEE
A51B	VAI SYLLABLE NJEE
A51C	VAI SYLLABLE YEE
A51D	VAI SYLLABLE KEE
A51E	VAI SYLLABLE NGGEE
A51F	VAI SYLLABLE GEE
A520	VAI SYLLABLE MEE
A521	VAI SYLLABLE NEE
A522	VAI SYLLABLE NYEE
A523	VAI SYLLABLE I
A524	VAI SYLLABLE HI
A525	VAI SYLLABLE HIN
A526	VAI SYLLABLE WI
A527	VAI SYLLABLE WIN
A528	VAI SYLLABLE PI
A529	VAI SYLLABLE BHI
A52A	VAI SYLLABLE BI
A52B	VAI SYLLABLE MBI
A52C	VAI SYLLABLE KPI
A52D	VAI SYLLABLE MGBI
A52E	VAI SYLLABLE GBI
A52F	VAI SYLLABLE FI
A530	VAI SYLLABLE VI
A531	VAI SYLLABLE TI
A532	VAI SYLLABLE THI
A533	VAI SYLLABLE DHI
A534	VAI SYLLABLE DHHI
A535	VAI SYLLABLE LI
A536	VAI SYLLABLE RI
A537	VAI SYLLABLE DI
A538	VAI SYLLABLE NDI
A539	VAI SYLLABLE SI
A53A	VAI SYLLABLE SHI
A53B	VAI SYLLABLE ZI
A53C	VAI SYLLABLE ZHI
A53D	VAI SYLLABLE CI
A53E	VAI SYLLABLE JI
A53F	VAI SYLLABLE NJI
A540	VAI SYLLABLE YI
A541	VAI SYLLABLE KI
A542	VAI SYLLABLE NGGI
A543	VAI SYLLABLE GI
A544	VAI SYLLABLE MI
A545	VAI SYLLABLE NI
A546	VAI SYLLABLE NYI
A547	VAI SYLLABLE A
A548	VAI SYLLABLE AN
A549	VAI SYLLABLE NGAN
A54A	VAI SYLLABLE HA
A54B	VAI SYLLABLE HAN
A54C	VAI SYLLABLE WA
A54D	VAI SYLLABLE WAN
A54E	VAI SYLLABLE PA
A54F	VAI SYLLABLE BHA
A550	VAI SYLLABLE BA
A551	VAI SYLLABLE MBA
A552	VAI SYLLABLE KPA
A553	VAI SYLLABLE KPAN
A554	VAI SYLLABLE MGBA
A555	VAI SYLLABLE GBA
A556	VAI SYLLABLE FA
A557	VAI SYLLABLE VA
A558	VAI SYLLABLE TA
A559	VAI SYLLABLE THA
A55A	VAI SYLLABLE DHA
A55B	VAI SYLLABLE DHHA
A55C	VAI SYLLABLE LA
A55D	VAI SYLLABLE RA
A55E	VAI SYLLABLE DA
A55F	VAI SYLLABLE NDA
A560	VAI SYLLABLE SA
A561	VAI SYLLABLE SHA
A562	VAI SYLLABLE ZA
A563	VAI SYLLABLE ZHA
A564	VAI SYLLABLE CA
A565	VAI SYLLABLE JA
A566	VAI SYLLABLE NJA
A567	VAI SYLLABLE YA
A568	VAI SYLLABLE KA
A569	VAI SYLLABLE KAN
A56A	VAI SYLLABLE NGGA
A56B	VAI SYLLABLE GA
A56C	VAI SYLLABLE MA
A56D	VAI SYLLABLE NA
A56E	VAI SYLLABLE NYA
A56F	VAI SYLLABLE OO
A570	VAI SYLLABLE HOO
A571	VAI SYLLABLE WOO
A572	VAI SYLLABLE WOON
A573	VAI SYLLABLE POO
A574	VAI SYLLABLE BHOO
A575	VAI SYLLABLE BOO
A576	VAI SYLLABLE MBOO
A577	VAI SYLLABLE KPOO
A578	VAI SYLLABLE MGBOO
A579	VAI SYLLABLE GBOO
A57A	VAI SYLLABLE FOO
A57B	VAI SYLLABLE VOO
A57C	VAI SYLLABLE TOO
A57D	VAI SYLLABLE THOO
A57E	VAI SYLLABLE DHOO
A57F	VAI SYLLABLE DHHOO

TABLE xx - Row A5: VAI

	A58	A59	A5A	A5B	A5C	A5D	A5E	A5F
0								
1								
2								
3								
4								
5								
6								
7								
8								
9								
A								
B								
C								
D								
E								
F								

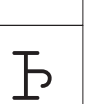
G = 00
P = 00

TABLE XX - Row A5: VAI

hex	Name
A580	VAI SYLLABLE LOO
A581	VAI SYLLABLE ROO
A582	VAI SYLLABLE DOO
A583	VAI SYLLABLE NDOO
A584	VAI SYLLABLE SOO
A585	VAI SYLLABLE SHOO
A586	VAI SYLLABLE ZOO
A587	VAI SYLLABLE ZHOO
A588	VAI SYLLABLE COO
A589	VAI SYLLABLE JOO
A58A	VAI SYLLABLE NJOO
A58B	VAI SYLLABLE YOO
A58C	VAI SYLLABLE KOO
A58D	VAI SYLLABLE NGGOO
A58E	VAI SYLLABLE GOO
A58F	VAI SYLLABLE MOO
A590	VAI SYLLABLE NOO
A591	VAI SYLLABLE NYOO
A592	VAI SYLLABLE U
A593	VAI SYLLABLE HU
A594	VAI SYLLABLE HUN
A595	VAI SYLLABLE WU
A596	VAI SYLLABLE WUN
A597	VAI SYLLABLE PU
A598	VAI SYLLABLE BHU
A599	VAI SYLLABLE BU
A59A	VAI SYLLABLE MBU
A59B	VAI SYLLABLE KPU
A59C	VAI SYLLABLE MGBU
A59D	VAI SYLLABLE GBU
A59E	VAI SYLLABLE FU
A59F	VAI SYLLABLE VU
A5A0	VAI SYLLABLE TU
A5A1	VAI SYLLABLE THU
A5A2	VAI SYLLABLE DHU
A5A3	VAI SYLLABLE DHHU
A5A4	VAI SYLLABLE LU
A5A5	VAI SYLLABLE RU
A5A6	VAI SYLLABLE DU
A5A7	VAI SYLLABLE NDU
A5A8	VAI SYLLABLE SU
A5A9	VAI SYLLABLE SHU
A5AA	VAI SYLLABLE ZU
A5AB	VAI SYLLABLE ZHU
A5AC	VAI SYLLABLE CU
A5AD	VAI SYLLABLE JU
A5AE	VAI SYLLABLE NJU
A5AF	VAI SYLLABLE YU
A5B0	VAI SYLLABLE KU
A5B1	VAI SYLLABLE NGGU
A5B2	VAI SYLLABLE GU
A5B3	VAI SYLLABLE MU
A5B4	VAI SYLLABLE NU
A5B5	VAI SYLLABLE NYU
A5B6	VAI SYLLABLE O
A5B7	VAI SYLLABLE ON
A5B8	VAI SYLLABLE NGON
A5B9	VAI SYLLABLE HO
A5BA	VAI SYLLABLE HON
A5BB	VAI SYLLABLE WO
A5BC	VAI SYLLABLE WON
A5BD	VAI SYLLABLE PO
A5BE	VAI SYLLABLE BHO
A5BF	VAI SYLLABLE BO
A5C0	VAI SYLLABLE MBO
A5C1	VAI SYLLABLE KPO
A5C2	VAI SYLLABLE MGBO
A5C3	VAI SYLLABLE GBO
A5C4	VAI SYLLABLE GBON
A5C5	VAI SYLLABLE FO
A5C6	VAI SYLLABLE VO
A5C7	VAI SYLLABLE TO
A5C8	VAI SYLLABLE THO
A5C9	VAI SYLLABLE DHO
A5CA	VAI SYLLABLE DHHO
A5CB	VAI SYLLABLE LO
A5CC	VAI SYLLABLE RO
A5CD	VAI SYLLABLE DO
A5CE	VAI SYLLABLE NDO
A5CF	VAI SYLLABLE SO
A5D0	VAI SYLLABLE SHO
A5D1	VAI SYLLABLE ZO
A5D2	VAI SYLLABLE ZHO
A5D3	VAI SYLLABLE CO
A5D4	VAI SYLLABLE JO
A5D5	VAI SYLLABLE NJO
A5D6	VAI SYLLABLE YO
A5D7	VAI SYLLABLE KO
A5D8	VAI SYLLABLE NGGO

hex	Name
A5D9	VAI SYLLABLE GO
A5DA	VAI SYLLABLE MO
A5DB	VAI SYLLABLE NO
A5DC	VAI SYLLABLE NYO
A5DD	VAI SYLLABLE E
A5DE	VAI SYLLABLE EN
A5DF	VAI SYLLABLE NGEN
A5E0	VAI SYLLABLE HE
A5E1	VAI SYLLABLE HEN
A5E2	VAI SYLLABLE WE
A5E3	VAI SYLLABLE WEN
A5E4	VAI SYLLABLE PE
A5E5	VAI SYLLABLE BHE
A5E6	VAI SYLLABLE BE
A5E7	VAI SYLLABLE MBE
A5E8	VAI SYLLABLE KPE
A5E9	VAI SYLLABLE KPEN
A5EA	VAI SYLLABLE MGBE
A5EB	VAI SYLLABLE GBE
A5EC	VAI SYLLABLE GBEN
A5ED	VAI SYLLABLE FE
A5EE	VAI SYLLABLE VE
A5EF	VAI SYLLABLE TE
A5F0	VAI SYLLABLE THE
A5F1	VAI SYLLABLE DHE
A5F2	VAI SYLLABLE DHHE
A5F3	VAI SYLLABLE LE
A5F4	VAI SYLLABLE RE
A5F5	VAI SYLLABLE DE
A5F6	VAI SYLLABLE NDE
A5F7	VAI SYLLABLE SE
A5F8	VAI SYLLABLE SHE
A5F9	VAI SYLLABLE ZE
A5FA	VAI SYLLABLE ZHE
A5FB	VAI SYLLABLE CE
A5FC	VAI SYLLABLE JE
A5FD	VAI SYLLABLE NJE
A5FE	VAI SYLLABLE YE
A5FF	VAI SYLLABLE KE

TABLE XX - Row A6: VAI

	A60	A61
0		
1		
2		
3		
4		
5		
6		
7		
8		
9		
A		
B		
C		
D		
E		
F		

G = 00
P = 00

TABLE XX - Row A6: VAI

hex	Name
A600	VAI SYLLABLE NGGE
A601	VAI SYLLABLE NGGEN
A602	VAI SYLLABLE GE
A603	VAI SYLLABLE GEN
A604	VAI SYLLABLE ME
A605	VAI SYLLABLE NE
A606	VAI SYLLABLE NYE
A607	VAI SYLLABLE NG
A608	VAI SYLLABLE LENGTHENER
A609	VAI COMMA
A60A	VAI FULL STOP
A60B	VAI QUESTION MARK
A60C	(This position shall not be used)
A60D	(This position shall not be used)
A60E	(This position shall not be used)
A60F	(This position shall not be used)
A610	VAI SYLLABLE NDOLE FA
A611	VAI SYLLABLE NDOLE KA
A612	VAI SYLLABLE NDOLE SOO
A613	VAI SYMBOL FEENG (thing)
A614	VAI SYMBOL KEENG (foot)
A615	VAI SYMBOL TING (island)
A616	VAI SYMBOL NII (cow)
A617	VAI SYMBOL BANG (finished)
A618	VAI SYMBOL FAA (die, kill)
A619	VAI SYMBOL TAA (go, carry, journey)
A61A	VAI SYMBOL DANG (hear, understand)
A61B	VAI SYMBOL DOONG (enter)
A61C	VAI SYMBOL KUNG (head, be able)
A61D	VAI SYMBOL TONG (be named)
A61E	VAI SYMBOL DO-O (be small)
A61F	VAI SYMBOL JONG (slave)

hex	Name

A. Administrative

1. Title

Proposal to add the Vai script to the BMP of the UCS.

2. Requester's name

Michael Everson, Charles Riley, José Rivera

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution.

4. Submission date

2005-08-01

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

Yes.

6b. More information will be provided later

No.

B. Technical -- General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

Yes.

Proposed name of script

Vai.

1b. The proposal is for addition of character(s) to an existing block

No.

1b. Name of the existing block

2. Number of characters in proposal

284.

3. Proposed category (see section II, Character Categories)

Category A.

4a. Proposed Level of Implementation (1, 2 or 3) (see clause 14, ISO/IEC 10646-1: 2000)

Level 1.

4b. Is a rationale provided for the choice?

Yes.

4c. If YES, reference

No combining characters.

5a. Is a repertoire including character names provided?

Yes.

5b. If YES, are the names in accordance with the character naming guidelines in Annex L of ISO/IEC 10646-1: 2000?

Yes.

5c. Are the character shapes attached in a legible form suitable for review?

Yes.

6a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Jason Glavy via Michael Everson. TrueType.

6b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson (Everson Typography). Fontographer.

7a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes (see above).

7b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

8. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

Yes. See above.

9. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script.

See above for character properties.

C. Technical -- Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

Yes. A preliminary discussion proposal was made in L2/05-053.

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?
Yes.
2b. If YES, with whom?
S. Jabaru Carlon (formerly University of Liberia), Mohamed B. Nyei (New York University), John V. Singler (New York University), Tombekai Sherman (Liberia), Jim Laesch (Lutheran Bible Translators), Miata Jessica Metzger (Cape Mount), Jumah Jennifer Gray Brumskine (Cape Mount), Lorna Priest (SIL), Peter Martin (SIL)
2c. If YES, available relevant documents
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?
See above.
4a. The context of use for the proposed characters (type of use; common or rare)
Used to write a the Mande language Vai used in Liberia.
4b. Reference
See above.
5a. Are the proposed characters in current use by the user community?
Yes.
5b. If YES, where?
Used in Liberia.
6a. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP?
Yes.
6b. If YES, is a rationale provided?
Yes.
6c. If YES, reference
Vai is a living script; BMP encoding is in accordance with the Roadmap.
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?
Yes.
8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?
No.
8b. If YES, is a rationale for its inclusion provided?
8c. If YES, reference
9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?
No.
9b. If YES, is a rationale for its inclusion provided?
9c. If YES, reference
10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?
No.
10b. If YES, is a rationale for its inclusion provided?
10c. If YES, reference
11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?
No.
11b. If YES, is a rationale for such use provided?
11c. If YES, reference
12a. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?
No.
12b. If YES, reference
13a. Does the proposal contain characters with any special properties such as control function or similar semantics?
No.
13b. If YES, describe in detail (include attachment if necessary)
14a. Does the proposal contain any Ideographic compatibility character(s)?
No.
14b. If YES, is the equivalent corresponding unified ideographic character(s) identified?
14c. If YES, reference